Progress and Prospects of U.S. Data Assimilation in Ocean Research

BY PIERRE F.J. LERMUSIAUX, PAOLA MALANOTTE-RIZZOLI, DETLEF STAMMER, JAMES CARTON, JAMES CUMMINGS, AND ANDREW M. MOORE

THIS REPORT summarizes goals, activities, and recommendations of a workshop on data assimilation held in Williamsburg, Virginia on September 9–11, 2003, and sponsored by the U.S. Office of Naval Research (ONR) and National Science Foundation (NSF). The overall goal of the workshop was to synthesize research directions for ocean data assimilation (DA) and outline efforts

Pierre F.J. Lermusiaux *(pierrel@pacific. deas.harvard.edu) is Research Associate, Harvard University, Cambridge, MA, USA.* **Paola Malanotte-Rizzoli** *is Professor of Physical Oceanography, Massachusetts Institute of Technology, Cambridge, MA, USA.* **Detlef Stammer** *is Professor of Physical Oceanography, Universität Hamburg, Germany.* **James Carton** *is Professor, Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA.* **James Cummings** *is Oceanographer, Naval Research Laboratory/Fleet Numerical Meteorology and Oceanography Center, Monterey, CA, USA.* **Andrew M. Moore** *is Professor, Department of Ocean Sciences, University of California, Santa Cruz, CA, USA.*

required during the next 10 years and beyond to evolve DA into an integral and sustained component of global, regional, and coastal ocean science and observing and prediction systems. The workshop built on the success of recent and existing DA activities such as those sponsored by the National Oceanographic Partnership Program (NOPP) and NSF-Information Technology Research (NSF-ITR).

DA is a quantitative approach to optimally combine models and observations. The combination is usually consistent with model and data uncertainties, which need to be represented. Ocean DA can extract maximum knowledge from the sparse and expensive measurements of the highly variable ocean dynamics. The ultimate goal is to better understand and predict these dynamics on multiple spatial and temporal scales, including interactions with other components of the climate system. There are many applications that involve DA or build on its results, including: coastal, regional, seasonal, and inter-annual ocean and climate dynamics; carbon and biogeochemical cycles; ecosystem dynamics; ocean engineering; observing-system design; coastal

management; fisheries; pollution control; naval operations; and defense and security. These applications have different requirements that lead to variations in the DA schemes utilized. For literature on DA, we refer to Ghil and Malanotte-Rizzoli (1991), the National Research Council (1991), Bennett (1992), Malanotte-Rizzoli (1996), Wunsch (1996), Robinson et al. (1998), Robinson and Lermusiaux (2002), and Kalnay (2003). We also refer to the U.S. Global Ocean Data Assimilation Experiment (GODAE) workshop on *Global Ocean Data Assimilation: Prospects and Strategies* (Rienecker et al., 2001); U.S. National Oceanic and Atmospheric Administration-Office of Global Programs (NOAA-OGP) workshop on *Coupled Data Assimilation* (Rienecker, 2003); and, NOAA-NASA-NSF workshop on *Ongoing Analysis of the Climate System* (Arkin et al., 2003).

WORKSHOP GOALS

To help the planning of sustained DA research, NSF and ONR requested a synthesis of current efforts and major research directions, including strategies for transitioning research results into

		Form Approved OMB No. 0704-0188

Report Documentation Page

1. REPORT DATE **MAR 2006**	2. REPORT TYPE	3. DATES COVERED **00-00-2006 to 00-00-2006**
4. TITLE AND SUBTITLE **Progress and Prospects of U.S. Data Assimilation in Ocean Research**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,Fleet Numerical Meteorology and Oceanography Center,Monterey,CA,93943**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**12**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

sustained activities that are useful for the whole community. Discussions were therefore centered on the needs of the community for all temporal and spatial scales; current status of ocean DA; major science issues and future challenges; transitions from the present status to sustained activities; and required funding and infrastructure issues. Both research and operational needs were addressed. The specific objectives of the workshop were to:

1. Assess the state of the art in DA and discuss the research required to realize its full potential.
2. Summarize the status of major national ocean DA activities and enumerate their impact.
3. Identify and review remaining challenges both in theory and applications.
4. Identify potential mechanisms for providing long-term funding support for sustained ocean DA.
5. Develop the interactions and transitions between research-oriented and operational efforts.
6. Discuss the needs for education and training and for retaining scientists involved in DA.

More details are provided at http://www.atmos.umd.edu/%7Ecarton/dameeting/abstracts.htm.

PURPOSES OF OCEAN DATA ASSIMILATION: FROM CLIMATE SCALES TO (SUB) MESOSCALES

Most DA purposes are common to all scales and applications. Purposes include: (1) to integrate the various data with dynamical models to obtain optimal descriptions of the ocean and allow accurate process studies, (2) to provide

a foundation for hypothesis testing and model improvement, including estimating model and data errors (uncertainty modeling), (3) to initialize ocean models, or the ocean component of coupled models, and assimilate subsequent observations for optimal forecasting, (4) to estimate model parameters and parameterizations, including forcing and lateral boundary conditions, and (5) to provide the means to assess observing systems, measure the utility of new data, and collect the most useful observations through adaptive sampling. With these purposes in mind, research applications are organized next by time scales, from the longest to the shortest.

Climate-System Time Scales

Due to its high capacity to store and redistribute heat, the ocean plays a central role in climate variability on time scales of tens to hundreds of years. To understand past and predict future climate changes, reasonable ocean estimates are necessary. Many key climate science questions (e.g., carbon cycling) involve interactions between the physical state of the ocean and its biogeochemistry. On global climatic scales, coupled oceanic and atmospheric processes are also important; it is crucial to refine and test air-sea fluxes to improve coupled models.

Realistic climate studies require a synthesis of all available and diverse data sets into a description of the four-dimensional evolution of the ocean consistent with a numerical dynamical model and its uncertainties. With such rigorous DA, momentum, heat, and salt are conserved. Major research issues in DA for climate purposes include:

(1) inherent biases in the model equations, numerical representations, multiple-scale data and observation operators; (2) errors of representativeness, or the errors in the data due to the presence of variability on scales distinct from climatic scales; (3) improvement of predictive skill by DA of mesoscale features into climate models; and (4) inference of sub-surface ocean fields from the four-dimensional DA of surface (satellite) information only.

Seasonal to Decadal Time Scales

At these scales, four processes have received special attention for fundamental and societal reasons. First, climate in many parts of the world is linked to the El Niño-Southern Oscillation (ENSO) in the tropical Pacific and Indian Oceans. Second, eastern North America and western Europe are influenced by a low-frequency oscillation in the latitude and intensity of Atlantic Ocean storms. Long multi-year fluctuations of this North Atlantic Oscillation (NAO) are linked to interactions between the atmosphere and the thermohaline circulation of the ocean. Third, the tropical Atlantic Ocean is subject to long multi-year variability (Tropical Atlantic Variability [TAV]) that depends on heat exchanges among the near-equatorial, tropical, and extra-tropical regions, and affects the eastern tropical Americas and western Africa. Fourth, the Pacific Decadal Oscillation (PDO) influences the climate of the North Pacific Ocean. It involves ocean-atmosphere interactions and changes in properties of major wind systems and of ocean variables such as sea surface temperature (SST).

Any comprehensive attempt to moni-

tor and eventually predict the above variability requires an accurate ocean estimate on at least basin scales. All DA approaches can be useful; DA already plays a key role for ENSO studies by providing initial conditions for quasi-operational climate forecasts. The four issues listed above for climate scales also apply to seasonal and decadal scales. Another issue directly relevant to decadal scales is the optimum reduction of the high dimensions of the DA problem so that advanced DA schemes can be used today in real time over ocean basins.

Monthly to Daily Time Scales

Mesoscale ocean dynamics—the internal weather of the sea—is highly energetic, variable, and eventful. The predictability of mesoscale ocean features such as fronts and eddies is limited because of their nonlinearities and chaotic nature. Although ocean mesoscale and atmospheric weather are similar, the ocean mesoscale has much shorter length scales. Considering characteristic spatial and temporal scales, ocean mesoscale predictions suffer from a lack of observations, particularly at depth. However, strong connections between large-scale and mesoscale dynamics allow the use of large-scale observation systems in mesoscale DA.

Ocean DA at monthly-to-daily scales requires an accurate forecast model and a DA scheme that best uses the sparse data. Simple schemes, such as optimal interpolation, have been very successful, though more advanced methods are needed. Major research issues in DA at these scales include: (1) optimal simplification of the problem so that advanced schemes can be used with comprehensive models at high resolution (10 km or less)

and over large domains (1000 km or more); (2) uncertainty predictions, including estimates of model and data errors, to improve models and observation systems; and (3) advanced algorithms (to reduce CPU time) and technology (distributed architectures).

Interdisciplinary Processes

Physical, biogeochemical, hydrological, and acoustic models are improving. The combination of disciplinary data and models allows the study and prediction of interdisciplinary interactions that occur in various ways, on multiple scales and with feedbacks. Generic scientific challenges include understanding regional and global biogeochemical cycles and marine ecosystems, the effects of human activities, and the interactions of various elements of the Earth system (e.g., hydrological cycles, geoid's undulations, glacial rebound). Novel utilizations of sound propagation for interdisciplinary monitoring are also emerging.

Although it is in its infancy, interdisciplinary DA can address the above challenges. Much can be accomplished now through existing methods, but care is needed due to nonlinearities, model errors, and data imprecision. Major interdisciplinary DA activities include: (1) carry out coupled acoustical-physical-biogeochemical-hydrological DA; (2) derive fundamental or simplified models and to estimate dominant term balances, adequate parameterizations, and proper aggregations of state variables; (3) model the sources of uncertainties in the multivariate data and models; and (4) estimate of the most useful data via adaptive sampling and observation-system simulation experiments.

PRESENT STATUS OF OCEAN DATA ASSIMILATION

The workshop presentations included presentations on DA systems and new projects. These presentations are summarized below.

DA Systems

The goal of the *HYbrid Cordinate Ocean Model* (HYCOM, which is part of GODAE) multi-institution partnership is to develop and apply eddy-resolving, real-time global and basin-scale prediction systems using HYCOM and to transition these systems to the U.S. Navy and NOAA (Chassignet et al., this issue). As part of HYCOM, a Hybrid Ocean Modeling Environment (HOME) will be developed and several ocean DA systems will be implemented and utilized. The first operational assimilation system will be a multivariate optimum interpolation (MVOI), with subsequent evaluations of the Reduced Order Information Filter (ROIF) and Singular Evolutive Extended Filter (SEEK). Atlantic and global configurations of the system will be used for varied applications. Products will be evaluated by comparison with independent data. The coastal ocean community is expected to use and evaluate products. Research directions will include correction of mean biases (e.g., sea surface height [SSH]), parameterizations (e.g., sea ice), grid resolutions, DA of new data types (e.g., SST, profiles, drifters), and product distribution.

NOPP established the *Estimating the Circulation and Climate of the Ocean* (ECCO) consortium (Figure 1) to bring ocean-state estimation to a quasi-operational tool for studying large-scale ocean dynamics, designing observational strate-
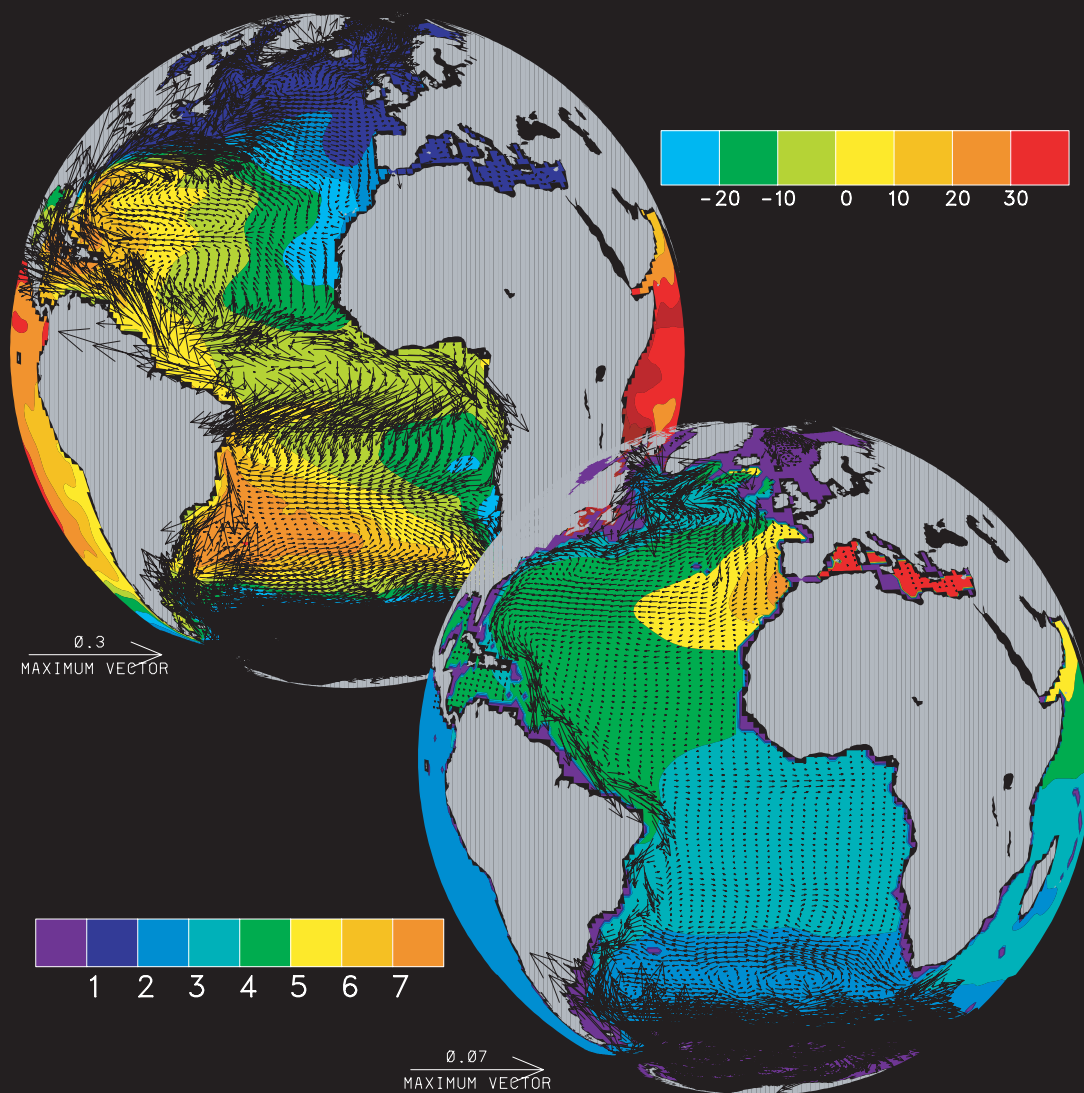
Figure 1. Mean flow field at 27 m and 1975 m depth from the 1 degree Estimating the Circulation and Climate of the Ocean (ECCO) synthesis calculation, together with the mean sea surface height (SSH) and the temperature field at 1975 m. The full time-dependent circulation is simulated by the ECCO effort since the beginning of 1992 in a way that is consistent with the surface forcing and the model dynamics. Although results are smooth due to the model resolution, the models transport estimates are optimally constrained by all available *in situ* and satellite data.

gies, and examining the ocean's role in climate variability. Important activities include re-analyses and initializations for forecasting. Model-data misfits are reduced based on adjoint assimilation and Kalman filtering and smoothing. The initial conditions and forcing are the control variables. Re-analysis fields for 11 years at one degree are available. A 50-year analysis that includes mixing as a control parameter is underway. The reduction of model-data misfits is estimated for each data type and the variability and mean of models and data are compared. Errors are small in the open ocean but are larger in western boundary regions. Research directions include comparisons to other products, time-dependent surface fluxes and transports, higher-resolutions including eddy-resolving DA, coupled and biogeochemical modeling, data and model error statistics, extended control space, parameterizations, observing system design, and sustained almost real-time global estimations.

The *Fleet Numerical Meteorology and Oceanography Center* (FNMOC) is one of the principal weather and ocean prediction centers within the U.S. Department of Defense (DOD). It runs a suite of global (NOGAPS, EFS, WW3, OTIS, TOPS) and regional (COAMPS, WW3, GFDN, PIPS) atmospheric and oceanographic modeling systems. Upcoming model improvements include increased resolutions, nesting, variational assimilation, boundary conditions, and new parameterizations. New parallel computing hardware (at the High-Performance

Sea Surface Height (cm)
NRL global NCOM glb8_2f
03-01-2006 00Z   0000 m

Sea Surface Temperature (C)
NRL global NCOM glb8_2f
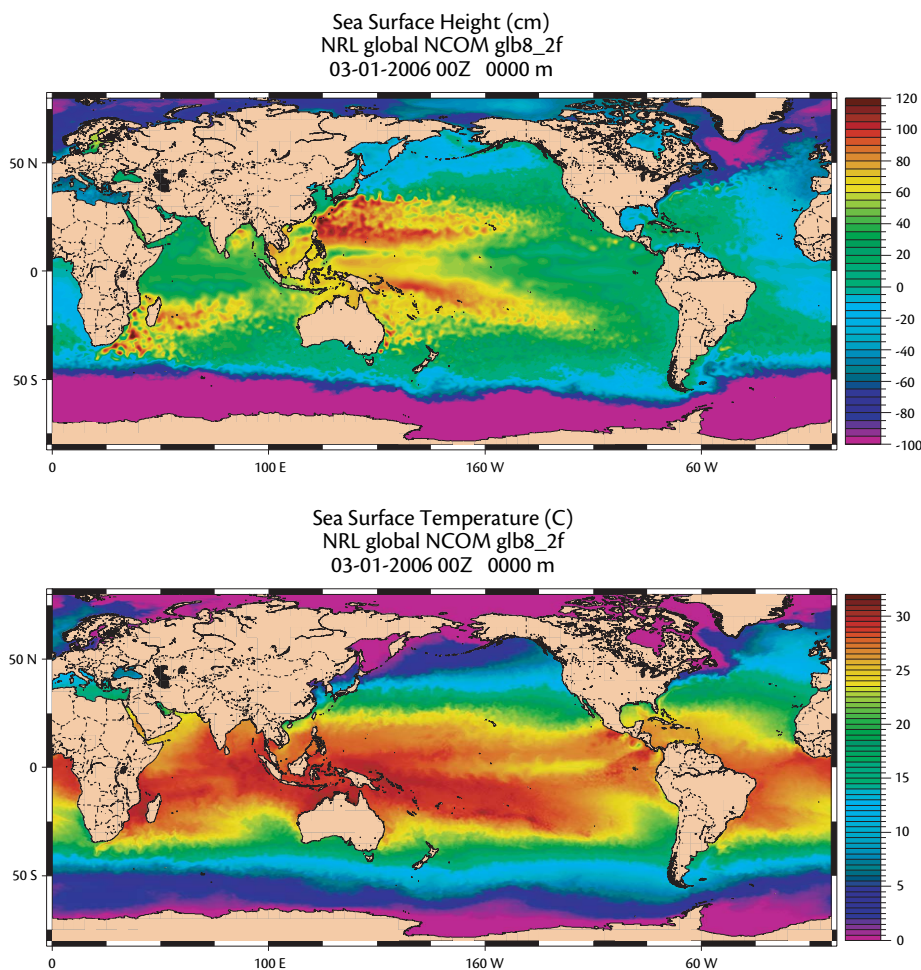## 03-01-2006 00Z   0000 m

Figure 2. Global NCOM SSH and SST analysis fields valid 1 March 2006 (C.N. Barron and J. Cummings, personal communication). Model resolution is 1/8th degree at mid-latitude. The global NCOM model assimilates altimeter SSH and satellite SST and is operational at NAVOCEANO. Global NCOM is the baseline system for all future Navy ocean DA developments (Barron et al., 2006).

Computer Modernization Program [HPCMP] center), GODAE data-server, ocean-atmosphere-ice coupling systems and new war-fighter products are in various levels of implementations.

The U.S. *Naval Oceanographic Office* (NAVOCEANO) (Figure 2) employs global (NLOM, G-NCOM), coastal (SWAFS, NCOM, WAM), bay (RMA-2, PC-Tides, etc) and dispersion (GNOME) models, with transitions from the Naval Research Laboratory (NRL) and an emphasis on rapid-response for coastal and bay users. The data assimilated include altimeter SSH, satellite SST, XBTs/CTDs, and tidal observations. NRL Layered Ocean Model (NLOM) (1/16th degree) tracks fronts and eddies with SSH/SST assimilation. NRL's Navy Coastal Ocean Model (NCOM)(1/8th degree) is based on the Princeton Ocean Model (POM) and provides boundary conditions for the Shallow Water Assimilation Forecast System (SWAFS) in nested domains. SWAFS is used for high-resolution coastal/bay simulations including drift predictions. RMA2 provides bays and estuary-river fields, with good operational success. The Modular Ocean DA System (MODAS) uses a T/S statistical model and optimal interpolation to provide gridded ocean fields. Research directions include efficient scheduling of (nested) jobs on parallel CPUs, higher resolutions, new parameterizations, and skill metrics relevant to the needs of the U.S. Navy.

The *Global Modeling and Assimilation Office* (GMAO) of the National Aeronautics and Space Administration/Goddard Space Flight Center (NASA/GSFC) results from the NASA Seasonal-to-Annual Prediction Project/Data Assimilation Office (NSIPP/DAO) merger. The thrust is to maximize the impact of satellite data in climate and weather predictions using global models and DA. Three assimilation systems are used: OI univariate, OI multivariate, and an ensemble Kalman filter. ENSO forecasts were improved by DA, which possibly indicates a breakdown of the spring predictability barrier. Better error covariance models led to better analysis. Coupled model and coupled initialization improved prediction skill. SSH assimilation enhanced forecasts in areas outside of the equatorial Pacific. Research directions and issues include observational errors, biases, model errors, and ensemble size and generation.

The *Upper Ocean Assimilation Systems* of the Geophysical Fluid Dynamics Laboratory/National Center for Environmental Prediction/Simple Ocean Data Assimilation system (GFDL/NCEP/SODA) (Figure 3) focuses on global

ocean DA, with an emphasis on the upper ocean. The NCEP effort is based on GFDL-MOM3 and 3D-VAR of XBTs, TAO moorings, and floats and altimeter data. Products include an ocean re-analysis from 1979 to the present, forced by daily NCEP fluxes. The GFDL effort is based on a quasi-global MOM4 model with 3D-VAR of XBT and SSH data, from 1993 to the present. SODA uses POP-2 on a global domain with sequential DA of a large set of hydrographic and satellite data, from 1940 up to today. It includes flow-dependent covariances and an explicit analysis of model bias. Comparisons among the three systems show that data coverage, data types assimilated, model resolutions, and parameterizations impact estimates, and that bias corrections (based on time-averaged data-model differences) improve forecasts substantially.

The *Systems Summary* reviewed why and how DA is carried out. Specific issues include multi-method approaches, systematic model errors, multivariate schemes, and data impacts. At monthly climate time scales, DA provides initial conditions for ocean-atmosphere forecasts and corrects for model errors. Adequate error metrics are very important. Because there is no unique purpose in DA, there may be no best method. ENhance ocean data Assimilation and ClimaTe (ENACT) is a multi-institution project to improve DA systems, produce global ocean analyses over multi-decadal periods and quantify enhancements through retrospective forecasts. This research addresses data impacts, use of different models and DA schemes, systematic errors, and measures of success. European collaborations (e.g., MERSEA,
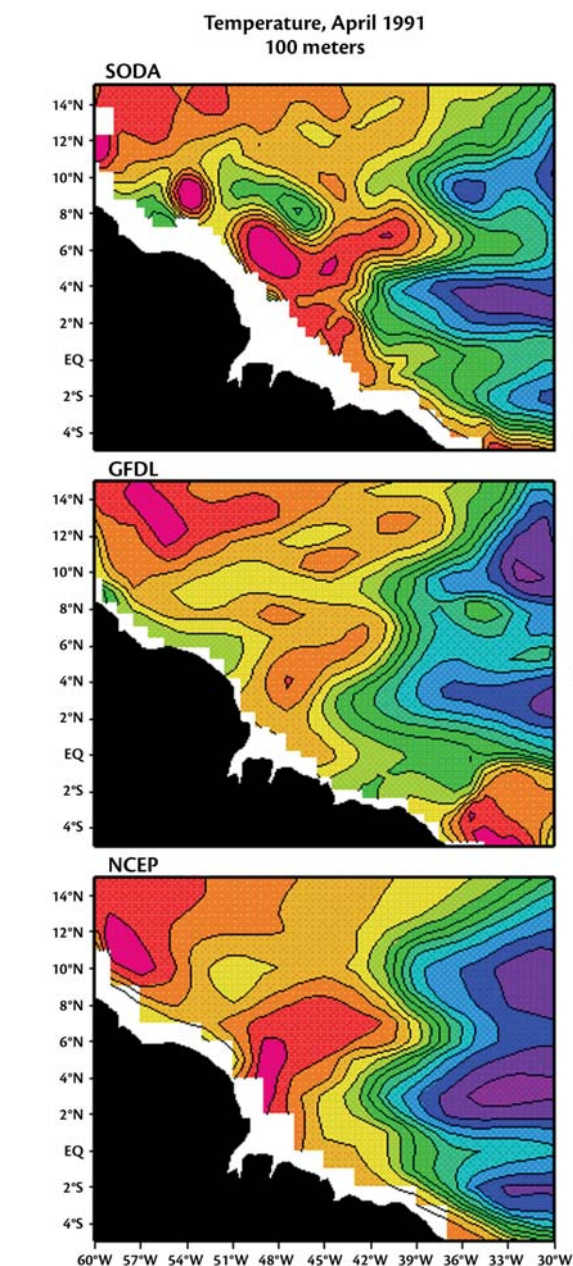

Temperature, April 1991
100 meters

Figure 3. Comparison of 100 m temperature from three re-analyses in the western tropical Atlantic for April, 1991. Temperature variability is introduced by northwestward moving warm core eddies containing water originating in the southern hemisphere. Differences in the representation of the eddies in three reanalyses, one from GFDL, one from NCEP, and one from University of Maryland (SODA, Carton and Giese, 2006) illustrate the challenge of producing accurate re-analyses of the mesoscale ocean prior to the launch of the *Topex/Poseiden* altimeters in 1992.

DUACS, DEMETER, ENSEMBLES) with ensembles of states, multi-model, multi-analyses, and multiple DA schemes have merit. Such U.S. collaborations should be useful and feasible. The examples illustrated show that ocean DA improves coupled forecasts, that adequate ensemble generations and models of model errors are essential, and that data-denial studies are needed. General research directions in ocean DA systems include systematic errors, bias corrections, and stochastic forcing; multivariate error covariances; coupled-models and multi-models; data impacts and sufficiency; effects of varied resolutions, scales, and processes; and multiple applications (e.g., coastal, biogeochemical, waves, sea-ice).

New Research Projects

The *Partnership for Advancing Interdisciplinary Global Modeling* (PARADIGM, NOPP) builds and deploys assimilative models of ecology and biogeochemistry for understanding and predicting ocean states in a range of environments (Rothstein et al., this issue). The scientific goal is to compare ecosystem dynamics in the North Pacific and Atlantic Ocean subtropical-subpolar gyres, including coastal regimes, with four themes: biogeochemical cycles, community structure, scales of physical forcing, and advanced interdisciplinary models. Technical goals are the merging of data and models, model (in)-validation, and statistical and dynamical system analyses. The physical models are ROMS, HYCOM, MIT-GCM,

and POP. Two DA objectives are to determine *a priori* what data resolution and types are needed and to simultaneously assimilate physical and biological data using adjoints and Kalman smoothers. Ongoing projects include constraining air-sea fluxes of carbon and oxygen with models and data, DA in the North Pacific, finding optical proxies of biological properties for DA in ecosystem models, developing new DA schemes, and studying the Northeast North America shelf carbon cycle.

The *Inverse Ocean Modeling System* (IOM, NSF- ITR) (Figure 4) is a multi-institution research project for a modular four-dimensional inversion system. The system inputs are the observing system, dynamics, hypotheses, estimator

and optimization algorithm. Its outputs are state estimates, data residuals, posterior statistics, model improvements and observing system assessments. The optimization algorithm is a four-dimensional variational least squares with weak constraints and inhomogeneous and colored covariances. The Euler-Lagrange equations are solved based on a representer approach. IOM uses a checklist on inputs and menu-like choices for the measurement model, dynamical model, forcing, initial/final and boundary conditions, parameters, and their respective prior errors. For example, if the dynamical model is nonlinear, the user provides the tangent-linear, representer and adjoint models. The outputs are user-dependent and the design is based on a user-friendly GUI-interface.

The *Littoral Ocean Observing and Prediction System* (*LOOPS*)/*Poseidon* (NSF-ITR) research (Figure 5) aims to develop a distributed system for real-time interdisciplinary ocean field and uncertainty forecasting with adaptive modeling and adaptive sampling. The scientific goal is to couple physical and biological oceanography with ocean acoustics. The technical goal is to build a dynamic system based on a web interface, metadata for model codes and data, distributed computing, legacy models and DA schemes, and new adaptive modeling and sampling. The legacy systems are the Harvard Ocean Prediction System (HOPS), Error Subspace Statistical Estimation (ESSE) and Massachusetts Institute of Technology (MIT) acoustical models. They are encapsulated at the binary level using XML and I/O workflows. ESSE is used for physical-acoustical and physical-biological DA. Adaptive
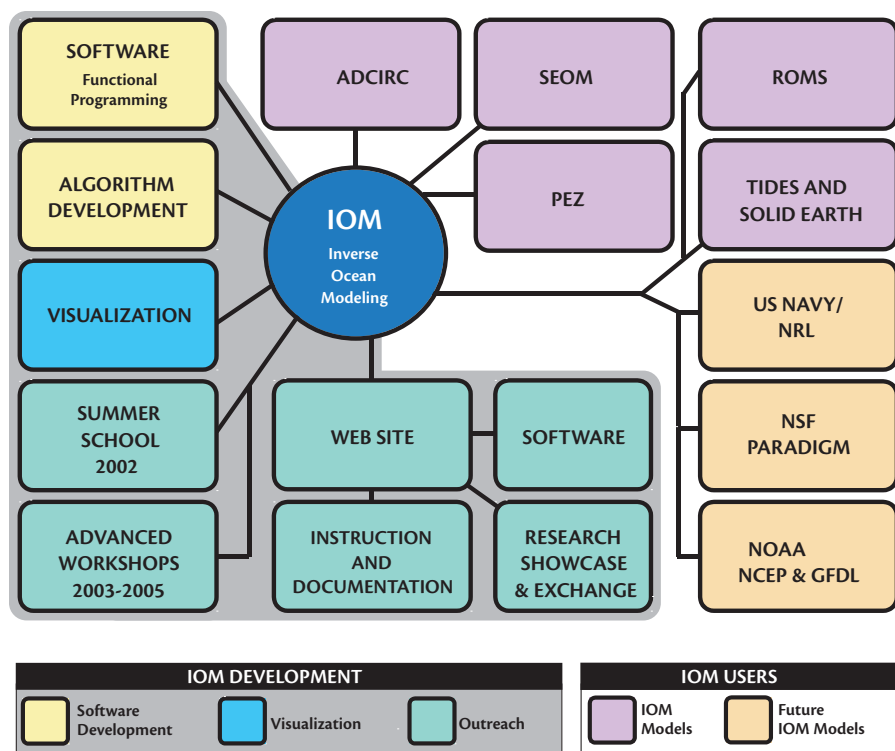

Figure 4. Organization chart of the Inverse Ocean Modeling (IOM) System. Courtesy of J. Muccino (Arizona State University) and A. Moore (University of California, Santa Cruz).

sampling forecasts the observing paths that minimizes uncertainties, optimizes the sampling of dynamical hot spots and maintains overall coverage. Adaptive modeling selects the physical or biological parameterizations that give the best model-data fit.

The *Environmental Data Assimilation Project* (NSF-ITR) (Figure 6) aims to investigate methodological issues that are common across Earth sciences (oceanography, meteorology, hydrology), with an emphasis on high dimensionality, nonlinearity, and uncertainties in models and measurements. Ensemble methods, and variational and automatic adjoint computations, are investigated. To account for uncertainties, it is natural to adopt a probabilistic approach, but di-

mensions are so large that conventional schemes cannot be used. In sequential estimation, three main approximations are possible. Quantization divides the ranges of random variables into finite discrete levels but the number of combinations remains too large. Particle filtering (Monte-Carlo) generates replicates of the state, uses the nonlinear model to evolve the ensemble, and updates replicates using Bayes theorem. It is not practical for oceanic problems, but can be used to test more practical ensemble schemes. The EnKF and related square-root schemes act as particle filtering, but use the sub-optimal minimum error variance update. Based on an idealized example from hydrology, it is shown that EnKF and ensemble square-root schemes

can be successful with highly non-Gaussian distributions, but that their posterior distributions can be too Gaussian.

The *Coastal Ocean Predictability and the Use of Ensemble Methods* was reviewed. Coastal primitive-equation models can now accurately represent and simulate the evolution of significant features of coastal ocean dynamics (e.g., wind responses, three-dimensional upwelling, wave-propagations, sea level, fronts, jets). However, several coastal processes (e.g., specific current instabilities and coupled air-sea interactions) are not yet predictable. Because the predictability of coastal winds is also limited, a statistical component should be introduced in dynamical predictions (i.e., ensemble forecasting). Ensembles can be generated from initial
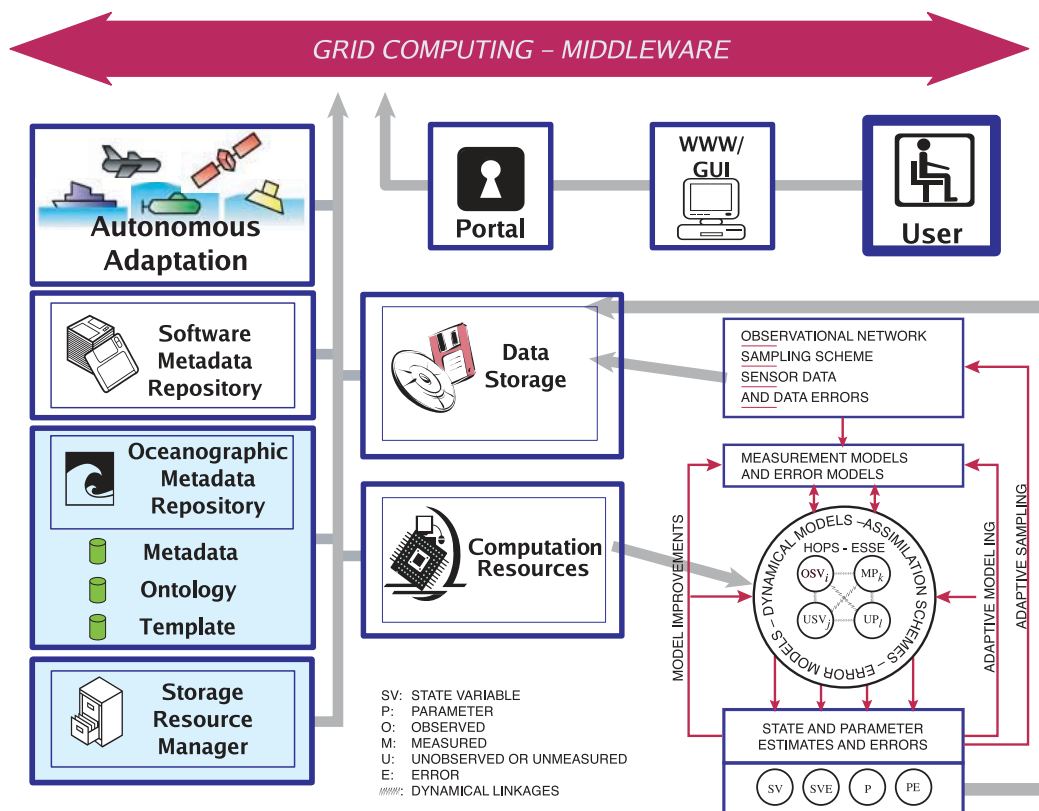

Figure 5. Schematic of the architecture of the Littoral Ocean Observing and Prediction System (LOOPS/*Poseidon*) (Patrikalakis et al, 2006). Remote compute and data resources, as well as adaptive observation systems, are connected to users through the use of grid computing middleware. Simulations and forecasts can be driven from a web browser and a portal gateway to the grid. Metadata for the software are written and utilized for web access. Ocean DA is carried out using the Harvard Ocean Prediction System (HOPS) and Error Subspace Statistical Estimation (ESSE). The repository for the metadata of ocean data and the manager of storage resources for keeping track of file locations are colored in light blue to indicate that they have not been developed.
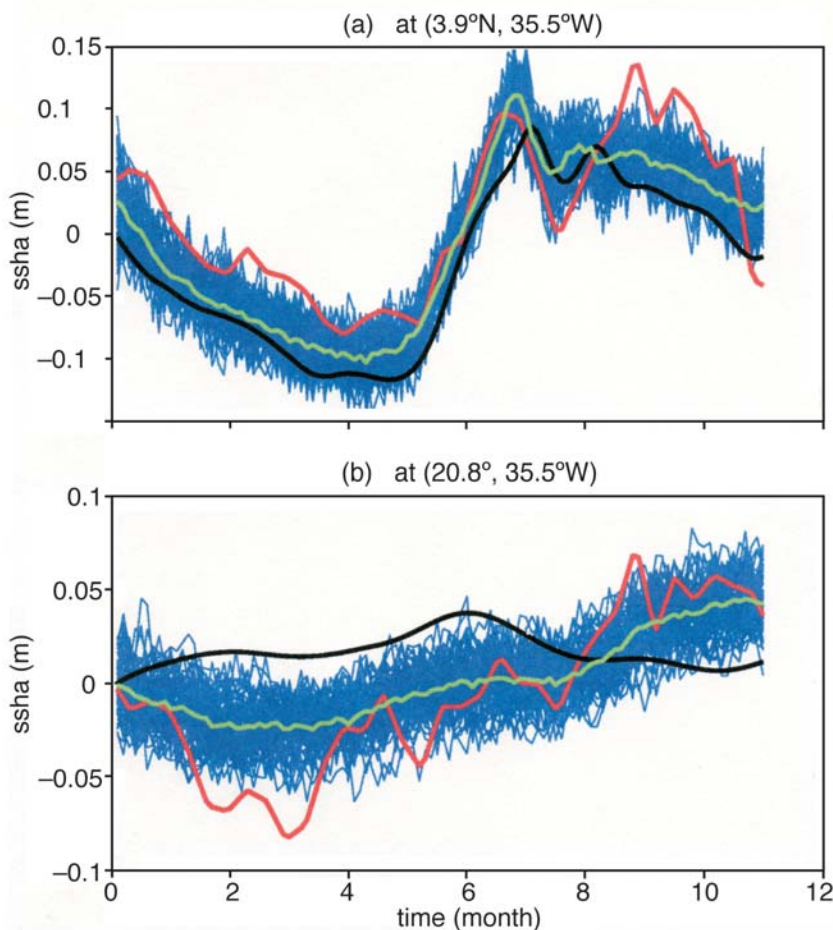
Figure 6. The Environmental Data Assimilation Project is illustrated by results of an estimation of the tropical Atlantic circulation from altimetry data using the ensemble Kalman filter (Zang and Malanotte-Rizzoli, submitted). Shown are the temporal evolution of sea surface height anomaly (ssha) anomaly from the observations (red line), the control run without DA (black line), and the mean of the analysis (green line) at two selected locations. The green lines are the ensemble analyses.

should be exploited. The future of all ocean-related work lies in the quantitative combination of models and data, but existing capabilities are at the margin of utility. Major technical issues include parameterizations or representations of the eddy-field, distinctions between systematic and stochastic errors, modeling and understanding of errors in very large systems, impacts of multiple scales on DA, and impacts of (open)-boundary conditions. A universally appropriate method is unlikely. Methodological issues include relations among the dimensions of the state vector, control vector, and ensemble; accuracy of error estimates; choice of norm or assimilation criterion; roles of global optimization schemes or explicitly Bayesian methods; and utility of Lagrangian approaches in very chaotic systems. For global estimations, maintaining and improving state estimation and observing systems on a global scale involves infrastructure issues with international implications. Global measurements will continue to be required in growing numbers and types. Other management issues include allowing non-experts to easily carry-out state estimation without misuse, maintaining an intellectual base in universities, and sustaining fundamental interest in large, complicated system research (e.g., Earth system), and providing action foci in institutions capable of supporting large applied engineering/numerical systems.

RECOMMENDATIONS AND SYNTHESIS OF PLENARY DISCUSSIONS

Plenary discussions and working groups identified major issues, emerging challenges, and future directions for ocean

conditions, model physics, and boundary forcing. Ensemble methods are informed by theories of disturbance growth because normal modes, Lyapunov vectors, and Floquet vectors are linked to bred vectors and singular vectors.

The *Challenges Summary* presentation reviewed where we want/need to be in 10–25 years, where the science and operational applications are going, and what the challenges are. Basic fluid-dy-namic understanding of the ocean relies on a combination of theory and data. Basic chemical/biological understanding of the ocean needs accurate fluid motions. Today, DA provides these motions, with a revolution in the making. Seasonal/inter-annual global forecasting is carried out, but the internal and coupled predictabilities need to be understood. Sensitivity studies based on historical climate records and adjoint solutions

DA research and operations. These discussions and reports were synthesized to lead the following set of major and immediate recommendations.

I. Ocean and Climate Sciences and Operations

Foster the scientific use of DA, with specific emphases on multiple scales, coupled ocean-atmosphere and land-surface processes, and interdisciplinary ocean science.

- **Multiple scales.** More efforts are needed on DA for the multiple and coupled oceanic scales, from the interactions of large-scales, mesoscales, and coastal scales, to turbulence effects in estuaries. Some of the generic research involves: (1) understanding multi-scale interactions, (2) inferring and modeling the effects of smaller scales in larger-scale simulations, (3) comparing the benefits of enhanced model resolution and model physics to those of enhanced estimation methods, (4) developing efficient DA schemes that explicitly address multiple scales and strong nonlinearities, and (5) developing multi-method DA approaches as well as DA schemes that learn from data. Examples of such research are quantifiying tidal-mesoscale interactions, finding adequate parameterizations of the eddy field in global and climate estimation, and nested modeling with scale-depend dynamics.

- **Oceanic-atmospheric-land processes.** Coupled DA will be essential for accurate simulations of such processes (e.g., CO_2 sequestration in the ocean). Generic research challenges include identifying the sources of errors in the coupled models and the data, and carrying-out efficient DA with comprehensive global models. Examples of the corresponding DA results would include climate forecasts with improved credibility.

- **Interdisciplinary processes.** DA efforts in the next decade and beyond will be multidisciplinary in nature as the separation between physical, acoustical, chemical, biogeochemical, and hydrological applications become transparent. Coupled interdisciplinary models and DA are a venue of the future. Some of the generic research involves: (1) using DA for fundamental research and process discovery in interdisciplinary ocean and climate science, (2) using DA to estimate the parameters and also the mathematical structures of biogeochemical models, (3) developing multi-method and adaptive DA for interdisciplinary studies, and (4) establishing the merits of nested or variable grids in ecosystem (e.g., for biological thin-layer dynamics) and acoustical modeling (e.g., for monitoring).

- **Ocean synthesis and reanalysis.** Ocean-state and parameter estimation via DA are carried out in several pilot settings, employing a variety of approaches and using the available data as constraints. Results will continue to be improved over the years to come, but many of such synthesis and reanalysis products are of sufficient quality to be used for scientific studies.

II. Hypotheses Testing and Uncertainty

Foster hypothesis-testing DA research to evaluate and improve models, estimate data and model uncertainties, and assess results.

- **Model testing (model-data comparisons).** Prior to DA, the underlying numerical model must be tested with respect to its skill in simulating observed ocean structures. This work is important for model improvement and is now an element of most large DA efforts. However, a detailed evaluation of the prior misfits in terms of model biases is outstanding, but could be very valuable.

- **Prior error estimates for models and data.** A necessary input for ocean DA is prior information about data and model uncertainties, including biases. Currently, neither of these error estimates exists with great detail. Much research is required in error modeling. The sources of model biases are usually hidden in forcing or in parameterizations. Tracing model biases back to these sources leads to corrections and understanding. Once errors are estimated and uncertainties modeled, they are input into the assimilation scheme. In general, uncertainty models should learn from the data and adapt to the model-data misfits. Importantly, external but also internal model errors, such as uncertainties on mixing coefficients or biological rates, should be included.

- **Uncertainty predictions and *a posteriori* error estimation.** Uncertainty prediction is challenging, in part due to limited computer resources. It is carried out today with reduced-dimension approaches and ensembles of simulations. Sequential estimation schemes need such uncertainty predictions for DA, but schemes based on adjoint equations often do not. However, all ocean DA systems can

and should provide some uncertainty estimates on the results, such as confidence intervals.

- **Assessment of DA results.** The *a posteriori* evaluation of error assumptions and residual statistics is important. It can reveal inconsistencies that have consequences for our understanding of ocean processes or numerical models. The evaluation will be most effective in multi-model, multi-method settings, so that the benefit of one or the other approach can be quantified. The tests should include a cross-validation of results against independent data and exploit predictions for validation.

III. Approaches and Methods

Sustain fundamental and applied research in the development of new DA schemes, including DA for model improvements and comparison and combination of DA methods.

- **DA for model and forcing improvements.** The success of DA depends on the quality of forward models and boundary conditions. Both have to be improved through repeated re-analyses and evaluations of data-model misfits. DA schemes should be extended to guide such improvements.
- **Multi-method and multi-model approaches.** Optimum DA methods have not been established for any major applications. In some cases, such as mesoscale or ENSO predictions, tuned suboptimal schemes might work or appear better than advanced dynamically consistent DA. A major recommendation is to combine multiple models and methods. By detailed comparisons of an ensemble of esti-

mates with each other and with ocean data, one can identify common (robust) results and separate them from model and/or method specific results. The differences discovered carry as much information as each solution. They allow estimating systematic model errors and the benefits or issues in individual methods.

- **New economic DA without lost of quality.** At present, DA efforts are fundamentally limited by the computational burden they entail. Promising directions include advanced adjoint-based systems, reduced-dimension schemes and ensemble-based methods, but improved methods are needed.

IV. Observations and Observing Systems

Bolster rigorous DA research and applications to optimize the design of observing systems, control the quality of measurements and adjust observational plans through adaptive sampling.

- **Observing system optimization.** In the past observing systems were set empirically. Through DA, one can now evaluate individual data sets and their role in constraining the estimate and helping our understanding of the ocean. Such optimization of the observing system using rigorous DA must be sustained. This allows revisiting the observing strategy and continuously adapting the network to the evolving dynamics as well as to new science questions.
- **Ocean data quality control.** Effective quality control of ocean data is essential. The use of model predictions as proxies in the quality control process is

promising. Comparing models to data in real-time as the data are collected can also help resolve issues of model skill, bias and lack of variability.

- **Adaptive sampling.** The path, locations and other properties of observing platforms and sensors can be optimized and adapted in real-time, so as to respond to the ocean variability and its uncertainties.

V. Technology, Systems, and Transfers

Link DA efforts to cyberinfrastructure initiatives and increase the support for sustained and efficient cooperative transitions of methodologies and technologies to operational centers.

- **Links to cyberinfrastructure.** Environmental cyberinfrastructure can be defined as a suite of critical enabling tools and research essential to study complex environmental ecosystems (NSF, 2003). Cyberinfrastructure (CI) provides tools for storing, finding, analyzing and synthesizing a diverse array of data. It also provides technologies for integrated multi-component modeling systems, visualizations, distributed computations, and adaptive workflows. In addition, CI supports the synthesis of observational data and models, provides collaboration tools, and offers new forms of education in science and engineering. It is of paramount importance to link now the development and sustained support of ocean DA research and systems to cyberinfrastructure initiatives.
- **Technology transfer.** An essential aspect of ocean DA research is the eventual transfer of technology to regional and national prediction centers. While

in principle such transfers are easy, they are challenging in practice. Issues range from lack of people and funding, to model incompatibilities and unmanageable computational burden. Even though some components of ocean DA, such as data quality control, data streams, and dissemination of results through common technology, are more easily transferable, all technology transfers require extensive and sustained cooperation.

VI. Required Infrastructures, Education and Funding

Create strong educational programs and career structures linked to ocean and climate DA, and provide a range of new funding opportunities for theoretical, applied, and operational efforts.

- **Human resources.** Only long-term funding in DA can establish adequate educational and career structures. Most needed items are: (1) sustained programs for students and research groups, (2) intellectual leadership by scientists in the growing DA community, and (3) career scientific programmers who are comfortable with evolving complex models, computer languages and cyberinfrastructure.

- **Computational support.** Support is required to sustain educational and operational facilities for ocean and climate DA. In practice, this implies new hardware acquisitions every three years. Computers don't need to be in the same physical location as users, but computer resources must be dedicated to the estimation activity. Because of ever-changing hardware, computer architectures, model codes, and standards, ocean and climate DA

activities must have ongoing cooperation with the computational sciences community (e.g., as in the NSF cyberinfrastructure projects).

- **Required funding.** We anticipate that NOPP will continue to be a source of funding for some DA activities, covering both personnel costs and hardware acquisitions. However, new other sustained sources of ocean DA support are urgently needed within the classic ONR and NSF funding programs, as well as within NOAA and NASA. Such support should involve a range of theoretical, applied, and operational possibilities, from small team efforts to larger-group collaborative research.

REFERENCES

Arkin, P., E. Kalnay, E. Laver, S. Schubert, and K. Trenberth. 2003. Report of the *Ongoing Analysis of the Climate System* workshop (NOAA-NASA-NSF). [Online] Available at: http://www.joss.ucar.edu/joss_psg/meetings/climatesystem/ [last accessed March 7, 2006].

Barron, C.N., A.B. Kara, P.J. Martin, R.C. Rhodes, and L.F. Smedstad. 2006. Formulation, implementation and examination of vertical coordinate choices in the global Navy Coastal Ocean Model (NCOM). *Ocean Modelling* 11(3–4):347–375, doi:10.1016/j.ocemod.2005.01.004.

Bennett, A.F. 1992. *Inverse Methods in Physical Oceanography*. Cambridge Monographs on Mechanics and Applied Mathematics. Cambridge University Press, Cambridge, United Kindgom.

Carton, J.A., and B.S. Giese. Submitted. SODA: A re-analysis of ocean climate, *Journal of Geophysical Research*.

Ghil, M., and P. Malanotte-Rizzoli. 1991. Data assimilation in meteorology and oceanography. In: *Advances in Geophysics*, 33, Academic Press, 141–266.

Kalnay, E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge, United Kingdom, 341 pp.

Malanotte-Rizzoli, P., ed. 1996. *Modern Approaches to Data Assimilation in Ocean Modeling*. Elsevier, Amsterdam, The Netherlands.

National Research Council (NRC). 1991. *Four-dimensional Model Assimilation of Data: A Strategy for the Earth System Sciences*. National Academy Press, Washington, D.C.

National Science Foundation (NSF). 2003. *Complex Environmental Systems: Synthesis for Earth, Life and Society in the 21st Century*. National Science Foundation, Advisory Committee on Environmental Research and Education, Arlington, VA, 80 pp. [Online] Available at: www.nsf.gov/geo/ere/ereweb/ac-ere/acere_synthesis_rpt_full.pdf [last accessed March 7, 2006].

Patrikalakis, N.M., J.J. McCarthy, A.R. Robinson, H. Schmidt, W. Cho, C. Evangelinos, P.J. Haley, S. Lalis, P.F.J. Lermusiaux, R. Tian, W.G. Leslie and W. Cho. Submitted. Towards a Dynamic Data-Driven System for Rapid Adaptive Interdisciplinary Ocean Forecasting.

Rienecker, M.M., D.E. Harrison, and W.B. Owens. 2001. Report of the US GODAE Workshop on *Global Ocean Data Assimilation: Prospects & Strategies*. [Online] Available at: http://www.usgodae.org/us-godae/wrkshp_apr_2001/usgodae_assim_report.pdf [last accessed March 7, 2006].

Rienecker, M.M. 2003. Report of the *Coupled Data Assimilation Workshop* (NOAA/OGP). [Online] Available at: www.amath.unc.edu/Faculty/nburrell/samsi/CoupledDA_rept_final.pdf [last accessed March 7, 2006].

Robinson, A.R., P.F.J. Lermusiaux, and N.Q. Sloan. 1998. Data Assimilation. In: *The Sea: The Global Coastal Ocean*, Vol. 10: Processes and Methods, K.H. Brink and A.R. Robinson, eds. John Wiley and Sons, New York, NY, 541–594.

Robinson, A.R., and P.F.J. Lermusiaux. 2002. Data assimilation for modeling and predicting coupled physical-biological interactions in the sea. In: *The Sea: Biological-Physical Interactions in the Sea*, Vol. 12, A.R. Robinson, J.J. McCarthy and B.J. Rothschild, eds. John Wiley and Sons, New York, NY, 475–536.

Wunsch, C. 1996. *The Ocean Circulation Inverse Problem*. Cambridge University Press, Cambridge, United Kingdom, 458 pp.

Zang, X. and P. Malanotte-Rizzoli. Submitted. Estimation of the tropical Atlantic circulation from altimetry data using the Ensemble Kalman Filter.